

Descritivo do Trabalho 1 sobre formação de Clusters

Exercícios sobre *Clustering* – Algoritmos de Agrupamento de Dados

Utilize o **Dataset dos Países** disponível no ambiente virtual. Utilize a biblioteca **ScikitLearn**.

Etapa Inicial (1 ponto) : selecione os atributos (colunas do dataset) que serão utilizados na formação de clusters.

Exercício 1 (2 pontos): Aplique o algoritmo **k-Means** nos dados do **Dataset** para encontrar 2 *clusters*.

- a) Descreva os dados que foram agrupados no *cluster* 0, usando o método de árvore de decisão (DecisionTreeClassifier).
- b) Descreva os dados que foram agrupados no *cluster* 1.
- c) Visualize os *clusters*, na formação dos dois *clusters*. Apresente os melhores resultados visualizados.

Exercício 2 (2 pontos): Aplique o algoritmo k-Means nos dados do **Dataset** para encontrar 3 *clusters*.

- a) Descreva os dados que foram agrupados no *cluster* 0.
- b) Descreva os dados que foram agrupados no *cluster* 1.
- c) Descreva os dados que foram agrupados no *cluster* 2.
- c) Visualize os *clusters*, na formação dos três *clusters*. Apresente os 3 melhores resultados visualizados.

Exercício 3 (1 pontos): Qual das duas formações de *clusters* é a melhor (exercício 1 ou 2)? Justifique a sua resposta.

Exercício 4 (4 pontos) : Realize testes com 4, 5, 6, etc *clusters*. Qual você considera o melhor parâmetro K, como resultado em termos de números de clusters, para o conjunto de dados fornecido? Apresente os clusters formados para o melhor valor K.

Entrega esperada: notebook do Colab contendo rótulos para identificação das respostas e códigos produzidos para cada uma dos exercícios.

